CLAIMS

1. A computer implemented method for gleaning file attributes independently of file format, the method comprising the steps of:

a non-application specific file attribute manager receiving a plurality of files in a plurality of formats;

the file attribute manager scanning the plurality of received files in the plurality of formats;

the file attribute manager gleaning attributes concerning each of the plurality of scanned files in the plurality of formats;

the file attribute manager storing gleaned attributes concerning each of the plurality of scanned files as records in a database; and

the file attribute manager indexing attributes being stored as a record in the database concerning a specific file according to contents of that file.

2. The method of claim 1 wherein:

the specific gleaned attributes concerning a specific file are a function of a protocol according to which the file is transmitted.

3. The method of claim 1 wherein:

the specific gleaned attributes concerning a specific file are a function of the format of that file.

4. The method of claim 1 further comprising:

the file attribute manager indexing attributes being stored as a record in the database concerning a specific file according to a secure hash of the contents of that file.

5. The method of claim 1 further comprising:

the file attribute manager indexing attributes being stored as a record in the database concerning a specific file according to a cyclical redundancy check of the contents of that file.

6. The method of claim 1 further comprising:

the file attribute manager receiving a plurality of copies of the same file; and the file attribute manager storing a separate record for each received copy of the file, each record being indexed according to the contents of the file, such that each record can be accessed by the single index.

7. The method of claim 1 further comprising:

deleting records from the database after the records have been stored for a specific period of time.

8. The method of claim 1 wherein the non-application specific file attribute manager is incorporated into at least one of the following:

a firewall;

an intrusion detection system;

an intrusion detection system application proxy;

a router;

a switch;

a standalone proxy;

a server;

a gateway;

an anti-virus detection system;

a client.

9. A computer readable medium containing a computer program product for gleaning file attributes independently of file format, the computer program product comprising program code for:

receiving a plurality of files in a plurality of formats;

scanning the plurality of received files in the plurality of formats;

gleaning attributes concerning each of the plurality of scanned files in the plurality of formats;

storing gleaned attributes concerning each of the plurality of scanned files as records in a database; and

indexing attributes being stored as a record in the database concerning a specific file according to contents of that file.

10. The computer program product of claim 9 further comprising:

program code for gleaning specific attributes concerning a specific file as a function of a protocol according to which the file is transmitted.

11. The computer program product of claim 9 further comprising:

program code for gleaning specific attributes concerning a specific file as a function of the format of that file.

12. The computer program product of claim 9 further comprising:

program code for indexing attributes being stored as a record in the database concerning a specific file according to a secure hash of the contents of that file.

13. The computer program product of claim 9 further comprising:

program code for indexing attributes being stored as a record in the database concerning a specific file according to a cyclical redundancy check of the contents of that file.

14. The computer program product of claim 9 further comprising:

program code for receiving a plurality of copies of the same file; and

program code for storing a separate record for each received copy of the file, each record being indexed according to the contents of the file, such that each record can be accessed by the single index.

15. The computer program product of claim 9 further comprising:

program code for deleting records from the database after the records have been stored for a specific period of time.

16. A computer system for gleaning file attributes independently of file format, the computer system comprising:

a reception module, configured to receive a plurality of files in a plurality of formats;

a scanning module, configured to scan the plurality of received files in the plurality of formats, the scanning module being communicatively coupled to the reception module;

a gleaning module, configured to glean attributes concerning each of the plurality of scanned files in the plurality of formats, the gleaning module being communicatively coupled to the scanning module;

a storage module, configured to store gleaned attributes concerning each of the plurality of scanned files as records in a database, the storage module being communicatively coupled to the gleaning module; and

an indexing module, configured to index attributes being stored as a record in the database concerning a specific file according to contents of that file, the indexing module being communicatively coupled to the storage module.

17. The computer system of claim 16 wherein:

the gleaning module is further configured to glean specific attributes concerning a specific file which are a function of a protocol according to which the file is transmitted.

18. The computer system of claim 16 wherein:

the gleaning module is further configured to glean specific attributes concerning a specific file which are a function of the format of that file.

19. The computer system of claim 16 wherein:
the indexing module is further configured to index attributes being stored as a record in the database concerning a specific file according to a secure hash of the contents of that file.

20. The computer system of claim 16 wherein:
the indexing module is further configured to index attributes being stored as a record in the database concerning a specific file according to a cyclical redundancy check of the contents of that file.

21. The computer system of claim 16 wherein:
the reception module is further configured to receive a plurality of copies of the same file; and
the storage module is further configured to store a separate record for each received copy of

the file, each record being indexed according to the contents of the file, such that each record can be accessed by the single index.

22. The computer system of claim 16 further comprising:

a deletion module, configured to delete records from the database after the records have been stored for a specific period of time, the deletion module being communicatively coupled to the storage module.

23. The method of claim 1 further comprising:

examining a file, the file having been processed by the non-application specific file attribute manager;

retrieving at least one stored record concerning the file from the database;

analyzing gleaned attributes concerning the file, the gleaned attributes having been retrieved from at least one record concerning the file in the database; and

responsive to analyzing the gleaned attributes, determining a status concerning the file.

24. The method of claim 23 further comprising: responsive to determining the status of the received file to be malicious, blocking the file.

25. The method of claim 23 further comprising: responsive to determining the status of the received file to be legitimate, not blocking the file.

26. The method of claim 23 further comprising: applying at least one rule specifying how to use gleaned file attributes to process the file.

27. The method of claim 26 further comprising: determining at least one of a plurality of rules to apply specifying how to use gleaned file attributes to process the file.